

ORIGINAL ARTICLE

Identifying newborns at risk of significant hyperbilirubinaemia: a comparison of two recommended approaches

R Keren, V K Bhutani, X Luan, S Nihtianova, A Cnaan, J S Schwartz

Arch Dis Child 2005;**90**:415–421. doi: 10.1136/adc.2004.060079

See end of article for authors' affiliations

Correspondence to:
Dr R Keren, The Children's
Hospital of Philadelphia,
3535 Market Street, Room
1524, Philadelphia, PA
19104, USA; keren@
email.chop.edu

Accepted
10 September 2004

Aims: To compare the predictive performance of clinical risk factor assessment and pre-discharge bilirubin measurement as screening tools for identifying infants at risk of developing significant neonatal hyperbilirubinaemia (post-discharge total serum bilirubin (TSB) >95th centile).

Methods: Retrospective cohort study of term and near term infants born in an urban community teaching hospital in Pennsylvania (1993–97). A clinical risk factor scoring system was developed and its predictive performance compared to a pre-discharge TSB expressed as a risk zone on a bilirubin nomogram. Main outcome measures were prediction model discrimination, range of predicted probabilities, and sensitivity, specificity, positive and negative predictive values, and likelihood ratios for various positivity criteria.

Results: The clinical risk factor scoring system developed included birth weight, gestational age <38 weeks, oxytocin use during delivery, vacuum extraction, breast feeding, and combination breast and bottle feeding. The pre-discharge bilirubin risk zone had better discrimination ($c=0.83$; 95% CI 0.80 to 0.86) than the clinical risk factor score ($c=0.71$; 95% CI 0.66 to 0.76) and predicted risk of significant hyperbilirubinaemia as high as 59% compared with a maximum of 44% for the clinical risk factor score. Neither the risk score nor the pre-discharge TSB risk zone predicted the outcome with ≥ 0.98 sensitivity without significantly compromising specificity (0.13 and 0.21, respectively). Multi-level clinical risk factor scores and TSB risk zones produced likelihood ratios of 0.15–3.25 and 0.05–9.43, respectively.

Conclusions: The pre-discharge bilirubin expressed as a risk zone on an hour specific bilirubin nomogram is more accurate and generates wider risk stratification than a clinical risk factor score.

Recognising the challenges of identifying infants with significant hyperbilirubinaemia after hospital discharge, the American Academy of Pediatrics (AAP) and other healthcare quality monitoring groups have recommended that prior to hospital discharge, providers screen all newborn infants for their risk of developing significant hyperbilirubinaemia. The two screening strategies that have been recommended are clinical risk factor assessment and bilirubin measurement prior to discharge.^{1–2} Precise risk factor assessment utilises clinical prediction rules that incorporate risk factors for hyperbilirubinaemia obtainable from the medical history and physical examination.³ The bilirubin measurement screening strategy utilises the infant's pre-discharge bilirubin expressed as a percentile on an hour specific bilirubin nomogram.⁴ The objective of this study was to assess and compare the predictive performance of these two screening strategies in order to identify a preferred clinical strategy.

METHODS

Design, setting, and study group

For the clinical risk factor screening strategy we developed a clinical risk factor based logistic regression model and then used that model to derive a clinical risk factor score to predict significant hyperbilirubinaemia. For the bilirubin measurement strategy, we evaluated the predictive accuracy of the pre-discharge bilirubin expressed as a risk zone on an hour specific bilirubin nomogram. The study group was drawn from infants who were the subjects of a previous study to develop the hour specific bilirubin nomogram.⁴ The base population was 13 003 full or near term infants born at an urban community teaching hospital in Pennsylvania during 1993–97. In order to exclude preterm infants with inaccurate

(overestimated) gestational age, the original study restricted subjects to infants with birth weight (BW) ≥ 2000 g if gestational age (GA) ≥ 36 weeks and BW ≥ 2500 g if GA ≥ 35 weeks. The study sample was further limited to infants who participated in the hospital's early discharge follow up programme ($n=2976$), which was offered to all mothers discharged within 48 hours after vaginal deliveries or 72 hours after caesarean sections who could not have their infant seen by a provider within a day or two. The early discharge follow up programme was introduced in 1993 in response to the relatively new but increasingly common practice of early postpartum discharge at that time. Infants who required phototherapy during the birth hospitalisation ($n=18$) or were treated in the intensive care nursery for any period of time prior to discharge ($n=118$) were excluded, leaving 2840 infants as potential subjects for analysis (fig 1).

As part of the early discharge follow up programme, an effort was made to obtain TSB levels in the hospital for all infants prior to discharge and on follow up. Throughout 1993–94, most infants had pre-discharge TSBs but only about 40% had both pre- and post-discharge TSBs. However, with the exception of 6 months, throughout 1995–97 more than 75% of all infants in the programme had both pre- and post-discharge TSB levels obtained. In order to prevent verification bias (which occurs when patients with a positive test result preferentially have the result confirmed with a reference standard test) we further restricted the study sample to infants born during these months in 1995–97 when $\geq 75\%$ of

Abbreviations: AAP, American Academy of Pediatrics; AGA, appropriate for gestational age; BW, birth weight; GA, gestational age; LGA, large for gestational age; SGA, small for gestational age; TSB, total serum bilirubin

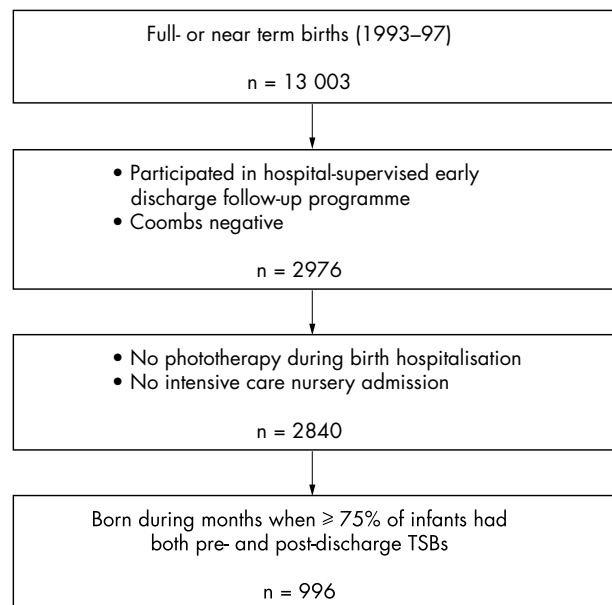


Figure 1 Assembly of study sample.

the infants had both pre- and post-discharge TSBs measured ($n = 996$) (fig 1). A total of 899 infants (90%) had both pre- and post-discharge TSBs in this time frame. The 97 infants who did not were not included in the development and evaluation of the risk assessment strategies. Institutional review boards of the participating institutions approved the study.

Data sources and collection

Demographic and laboratory data were obtained from the research database developed for the original bilirubin nomogram study. Research assistants reviewed the original maternal and infant birth hospitalisation records of all 2840 infants from that study to collect additional data about clinical risk factors not recorded as part of that study. Additional demographic information (such as maternal race and age) was abstracted from standard administrative forms. Maternal and infant clinical information, such as medical history and physical findings, was drawn primarily from standard admission, intrapartum, and discharge forms. When clinical information on these forms was either missing or ambiguous, research assistants reviewed physician progress notes and nursing flow sheets for additional information. Two research assistants performed the chart reviews and entered data into a Microsoft Access (Redmond, WA) relational database.

Data elements and definitions

The dependent variable for the prediction models was development of significant hyperbilirubinaemia, defined as a post-discharge TSB >95th centile on an hour specific bilirubin nomogram.⁴ The 95th centile on the bilirubin nomogram is nearly identical to the phototherapy threshold curve recommended for “medium risk” infants in the AAP’s 2004 clinical practice guideline, and thus is a clinically meaningful outcome to predict.²

We identified risk factors from the medical history, physical examination, or laboratory evaluation that were previously reported as predictors of significant neonatal hyperbilirubinaemia.^{3–5–8} These are grouped into infant factors, maternal factors, and pregnancy events/delivery characteristics in table 1. Because the presence and degree of jaundice often cannot be accurately nor reliably estimated⁹

and because it was not systematically or consistently documented in the hospital record, we did not include it as a predictor in the development of the clinical risk factor score.

Pre/post discharge TSB and risk zone designation

When an infant had more than one pre-discharge TSB, the TSB measurement closest to the time of hospital discharge was selected as the predictor TSB. In general, TSBs obtained prior to and after hospital discharge were defined as pre- and post-discharge TSBs, respectively. However, to approximate commonly occurring newborn lengths of stay, for a minority of cases we used two rules developed a priori to reassign pre-discharge TSB to post-discharge TSB, and vice-versa. First, when an infant remained in the hospital for more than two days (for example, born by caesarean section) and had multiple TSB values obtained prior to discharge, but none after discharge, if the last TSB value was obtained after 40 hours of age, then the TSBs obtained prior to 40 hours were designated pre-discharge TSBs and the subsequent TSBs were designated post-discharge TSBs. Second, when an infant was discharged early and had no pre-discharge TSB but did have multiple post-discharge TSBs, if there was at least one TSB obtained before 72 hours, then the first TSB was designated a pre-discharge TSB and all subsequent TSBs were defined as post-discharge TSBs. Overall, we reassigned a pre-discharge TSB to a post-discharge TSB for 213 (7.5%) infants and a post-discharge TSB to a pre-discharge TSB for 96 (3.4%) infants.

Analysis

Univariate associations between the individual predictors and the development of a post-discharge hour specific TSB >95th centile, were examined using t tests for continuous predictors and Pearson χ^2 and Fisher’s exact test (when appropriate) for categorical predictors. Clinical risk factors (not including the pre-discharge TSB or risk zone) that were associated with the outcome in univariate analyses at the $p < 0.2$ level of significance were considered for inclusion in a clinical risk factor based multivariable logistic regression model. In cases where the candidate variables were highly correlated (for example, birth weight and discharge weight), the most clinically relevant and reliably ascertained variable was chosen for inclusion in the model building procedure. The multivariable model was developed using the best subset selection method^{10–11} and compared to those developed using forward selection, backward elimination, and stepwise procedures. To translate the final logistic regression model into a clinical risk factor score, each infant was assigned points, equal to the sum of the odds ratios (ORs) (multiplied by 2 and rounded down) corresponding to risk factors present for that infant.³ Scoring systems based on the logistic coefficients and non-multiplied ORs produced identical results but involved adding fractions, which we felt was cumbersome for clinicians.

The hour specific bilirubin nomogram was used to translate the pre-discharge predictor TSB values into corresponding centile based risk zones. The pre-discharge risk zone (low, 0–40th centile; low-intermediate, 41th–75th centile; high-intermediate, 76th–95th centile; and high, >95th centile) was then used to predict the outcome of interest.

The accuracy of the clinical risk factor score and the pre-discharge TSB risk zone as predictors of subsequent significant hyperbilirubinaemia was compared in three ways. First we used the c-statistic¹² to calculate model discrimination, that is, the models’ ability to distinguish between patients with and without the outcome of interest (development of significant hyperbilirubinaemia). The c-statistic is equivalent to the area under the receiver operating characteristics

Table 1 Relation of infant, maternal, and pregnancy/delivery characteristics to the development of a post-discharge TSB >95th centile*

| Variables | Overall | Post-discharge TSB >95th centile | | p value | OR (95% CI) |
|---|------------|----------------------------------|--------------|---------|---------------------|
| | | No (n = 801) | Yes (n = 98) | | |
| Infant | | | | | |
| Male (no., %) | 471 (52) | 418 (52) | 53 (54) | 0.72 | 1.1 (0.7 to 1.6) |
| Cephalohaematoma (no., %) | 55 (6) | 49 (6) | 6 (6) | 0.97 | 1.0 (0.4 to 2.4) |
| Bruising (no., %) | 41 (5) | 36 (4) | 5 (5) | 0.80 | 1.2 (0.4 to 3.0) |
| Gestational age (no., %) | | | | | |
| <38 wk | 179 (20) | 152 (19) | 27 (28) | 0.02 | 1.9 (1.1 to 3.1) |
| 38–39 wk† | 493 (55) | 450 (56) | 43 (44) | – | – |
| ≥40 wk | 227 (25) | 199 (25) | 28 (29) | 0.13 | 1.5 (0.9 to 2.4) |
| Size for gestational age (no., %) | | | | | |
| SGA | 69 (8) | 68 (9) | 1 (1) | 0.04 | 0.1 (0.02 to 0.9) |
| AGA† | 739 (82) | 657 (82) | 82 (84) | – | – |
| LGA | 91 (10) | 76 (10) | 15 (15) | 0.13 | 1.6 (0.9 to 2.9) |
| Pre-discharge TSB (no., %) | | | | | |
| <40th centile† | 172 (19) | 171 (21) | 1 (1) | – | – |
| 40th–75th centile | 398 (44) | 388 (48) | 10 (10) | 0.16 | 4.4 (0.6 to 35) |
| 76th–95th centile | 273 (30) | 216 (27) | 57 (58) | <0.001 | 45 (6.2 to 329) |
| >95th centile | 56 (6) | 26 (3) | 30 (31) | <0.001 | 197 (26 to 1509) |
| Birth weight, kg (mean (SD)) | 3.3 (0.5) | 3.3 (0.5) | 3.4 (0.5) | 0.008 | 1.8 (1.2 to 2.8) |
| % weight loss/day (mean (SD)) | 2.1 (1.6) | 2.1 (1.5) | 2.1 (2.4) | 0.97 | 1.0 (0.9 to 1.1) |
| Maternal | | | | | |
| Race/ethnicity (no., %) | | | | | |
| White† | 390 (43) | 345 (43) | 45 (46) | – | – |
| Black | 351 (39) | 320 (40) | 31 (32) | 0.23 | 0.7 (0.5 to 1.2) |
| Hispanic | 39 (4) | 36 (5) | 3 (3) | 0.47 | 0.6 (0.2 to 2.2) |
| Asian | 72 (8) | 62 (8) | 10 (10) | 0.57 | 1.2 (0.6 to 2.6) |
| Others | 46 (5) | 38 (5) | 8 (8) | 0.25 | 1.6 (0.7 to 3.7) |
| Method of feeding | | | | | |
| Breast | 461 (51) | 397 (50) | 64 (65) | <0.001 | 2.5 (1.5 to 4.2) |
| Bottle† | 332 (37) | 312 (39) | 20 (20) | – | – |
| Both | 104 (12) | 90 (11) | 14 (14) | 0.02 | 2.4 (1.2 to 5.0) |
| Hypertension (no., %) | 19 (2) | 16 (2) | 3 (3) | 0.44 | 1.6 (0.5 to 5.6) |
| Diabetes mellitus (no., %) | 24 (3) | 19 (3) | 5 (6) | 0.17 | 2.3 (0.8 to 6.2) |
| Age (mean (SD)) | 28.5 (6.3) | 28.4 (6.3) | 29.2 (6.1) | 0.22 | 1.02 (0.99 to 1.06) |
| Parity (mean (SD)) | 1.0 (1.2) | 1.0 (1.3) | 0.7 (1.0) | 0.03 | 0.8 (0.6 to 0.98) |
| Pregnancy and delivery | | | | | |
| Vaginal bleeding (no., %) | 3 (0.4) | 3 (0.4) | 0 (0) | 0.98 | 0 |
| Smoked (no., %) | 69 (8) | 63 (8) | 6 (6) | 0.53 | 0.8 (0.3 to 1.8) |
| Caesarean section (no., %) | 101 (11) | 94 (12) | 7 (7) | 0.18 | 0.6 (0.3 to 1.3) |
| Forceps (no., %) | 42 (5) | 39 (5) | 3 (3) | 0.61 | 0.6 (0.2 to 2.1) |
| Vacuum (no., %) | 133 (15) | 107 (13) | 26 (27) | <0.001 | 2.4 (1.4 to 3.9) |
| Prolonged rupture of membranes (no., %) | 35 (4) | 28 (4) | 7 (7) | 0.08 | 2.1 (0.9 to 5.0) |
| Oxytocin (no., %) | 542 (61) | 469 (60) | 73 (75) | 0.002 | 2.1 (1.3 to 3.4) |
| Epidural (no., %) | 635 (71) | 562 (71) | 73 (75) | 0.35 | 1.3 (0.8 to 2.1) |

*Gestational age based on Ballard score and obstetrical dating of pregnancy rounded to the nearest week; small for gestational age (SGA) defined as birth weight for gestational age below the 10th centile and large for gestational age (LGA) as at or above the 90th centile (stratified by sex and race);²⁶ maternal race was self-reported at the time of registration and documented in administrative records; planned method of feeding was reported by mothers to nurses at discharge and documented in discharge instruction forms; gestational diabetes diagnosis based on glucose tolerance test (GTT) results; parity figure does not include infant delivered at birth hospitalisation; vaginal bleeding refers to bleeding occurring after 20 weeks of pregnancy; prolonged rupture of membranes defined as greater than 18 hours; use of oxytocin for induction or augmentation (but not post-partum) defines oxytocin exposure.

†Reference group for OR and p value calculations.

Table 2 Variables associated with post-discharge TSB >95th centile in clinical risk factor logistic regression model and corresponding clinical risk factor score

| Variable | OR (95% CI) | p value | Clinical risk factor score* |
|---------------------------|------------------|---------|-----------------------------|
| Birth weight† | 1.5 (1.2 to 1.9) | 0.001 | |
| 2000–2500 g | | | 0 |
| 2501–3000 g | | | 3 |
| 3001–3500 g | | | 6 |
| 3501–4000 g | | | 9 |
| 4001–4500 g | | | 12 |
| 4501–5000 g | | | 15 |
| GA<38 weeks | 2.6 (1.5 to 4.5) | 0.001 | 5 |
| Oxytocin | 2.0 (1.2 to 3.4) | 0.005 | 4 |
| Vacuum delivery | 2.2 (1.5 to 3.6) | 0.003 | 4 |
| Exclusive breast feeding | 2.6 (1.5 to 4.5) | <0.001 | 5 |
| Breast and bottle feeding | 2.3 (1.1 to 4.9) | 0.03 | 4 |

*Points for each clinical risk factor were generated by multiplying the corresponding OR by two and rounding down to the nearest integer.

†Infant birth weight (kg) was transformed by subtracting 2 kg and dividing by 0.5 in order to generate the odds ratio for increased risk of significant hyperbilirubinaemia for every 0.5 kg birth weight above 2.5 kg.

(ROC) curve,¹³ which we plotted for each model. The discrimination of the models was compared using an algorithm developed by DeLong and colleagues.¹⁴ Second, we determined and compared the spectrum of predicted probabilities generated by each model. Third, we calculated the sensitivity, specificity, positive and negative predictive values, and likelihood ratios for alternative positivity criteria for each prediction rule. The Hosmer-Lemeshow goodness-of-fit statistic¹⁵ was used to determine each prediction rule's calibration, that is, the degree of correlation between observed and predicted rates of the outcome along the spectrum of predicted risk.

RESULTS

Approximately 11% (n = 98) of infants developed significant hyperbilirubinaemia (a post-discharge TSB >95th centile). In univariate analyses, factors that were associated with an increased risk of developing significant hyperbilirubinaemia (p < 0.20) included gestational age (GA) <38 weeks and ≥40 weeks, large for gestational age (LGA), higher pre-discharge TSB risk zone, higher birth weight, breast feeding, combined breast and bottle feeding, maternal diabetes, vacuum extraction, prolonged rupture of membranes, and oxytocin use. Small for gestational age (SGA), parity, and caesarean section were associated with a decreased risk of developing significant hyperbilirubinaemia (table 1).

Clinical risk factor score

The clinical risk factor scoring system derived from the multivariable logistic regression included the following risk factors: birth weight, GA <38 weeks, oxytocin use during delivery, vacuum extraction, breast feeding, and combination breast and bottle feeding (table 2). Model building using forward selection, backward elimination, and stepwise procedures identified the same risk factors, except the combined breast and bottle feeding variable did not remain in the models. The clinical risk factor score had a c-statistic of 0.71 (95% CI 0.66 to 0.76) (fig 2). The range of predicted probabilities generated by the scoring system was relatively narrow (0.01–0.44).

Table 3 shows the predictive properties of the clinical risk factor score using various cut-off points for what score is considered a positive test. If the goal of screening is to identify all true positives, then a relatively low score needs to be used as the positivity criterion. For example, using a score of 8 as the positivity criterion identified 98% of infants who went on to develop significant hyperbilirubinaemia (that is,

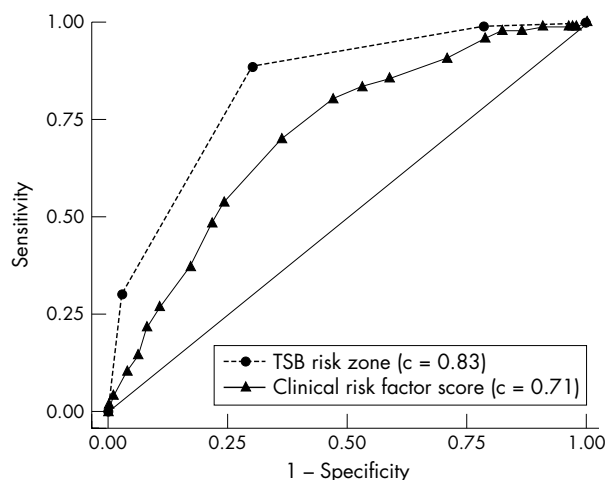


Figure 2 Receiver operating curves for clinical risk factor score and pre-discharge TSB risk zone.

Table 3 Predictive properties of the clinical risk factor score and pre-discharge TSB risk zone using various cut-off points for a positive test

| Outcome | | Predictive properties | | | | | | |
|--|---|-----------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|----------------------|---------------------|
| Positive test (no.) | Post-discharge TSB > 95th centile (no.) | Sensitivity (95% CI) | | | | | | |
| | | Specificity (95% CI) | Positive predictive value (95% CI) | Negative predictive value (95% CI) | Likelihood ratio positive (95% CI) | Likelihood ratio negative (95% CI) | | |
| Positivity criterion (risk factor score) | | | | | | | | |
| ≥0 | 884* | 97 | 1.0 (0.96 to 1.0) | 0 (0 to 0) | 0.11 (0.09 to 0.13) | – | 1.0 (1.0 to 1.0) | – |
| ≥8 | 777 | 95 | 0.98 (0.93 to 1.0) | 0.13 (0.11 to 0.16) | 0.12 (0.10 to 0.15) | 0.98 (0.93 to 1.0) | 1.13 (1.13 to 1.13) | 0.15 (0.06 to 0.41) |
| ≥12 | 547 | 83 | 0.86 (0.77 to 0.92) | 0.41 (0.38 to 0.45) | 0.15 (0.12 to 0.18) | 0.96 (0.93 to 0.98) | 1.45 (1.44 to 1.47) | 0.35 (0.31 to 0.40) |
| ≥16 | 244 | 52 | 0.54 (0.43 to 0.64) | 0.76 (0.72 to 0.79) | 0.21 (0.16 to 0.27) | 0.93 (0.91 to 0.95) | 2.20 (2.15 to 2.26) | 0.61 (0.59 to 0.63) |
| ≥20 | 86 | 21 | 0.22 (0.14 to 0.31) | 0.92 (0.90 to 0.94) | 0.24 (0.16 to 0.35) | 0.90 (0.88 to 0.92) | 2.62 (2.45 to 2.89) | 0.85 (0.84 to 0.86) |
| ≥24 | 7 | 2 | 0.02 (0 to 0.07) | 0.99 (0.99 to 1.00) | 0.29 (0.04 to 0.71) | 0.89 (0.87 to 0.91) | 3.25 (1.50 to 10.44) | 0.99 (0.98 to 0.99) |
| Positivity criterion (centile) | | | | | | | | |
| ≥0th | 899 | 98 | 1.0 (0.96 to 1.0) | 0 (0.00 to 0.00) | 0.11 (0.09 to 0.13) | – | 1.0 (1.0 to 1.0) | – |
| ≥40th | 727 | 97 | 0.99 (0.94 to 1.00) | 0.21 (0.19 to 0.24) | 0.13 (0.11 to 0.16) | 0.99 (0.97 to 1.0) | 1.26 (1.26 to 1.26) | 0.05 (0.01 to 0.34) |
| ≥75th | 329 | 87 | 0.89 (0.81 to 0.94) | 0.70 (0.66 to 0.73) | 0.26 (0.22 to 0.32) | 0.98 (0.97 to 0.99) | 2.94 (2.92 to 2.97) | 0.16 (0.14 to 0.19) |
| ≥95th | 56 | 30 | 0.31 (0.22 to 0.41) | 0.97 (0.95 to 0.98) | 0.54 (0.40 to 0.67) | 0.92 (0.90 to 0.94) | 9.43 (9.27 to 10.34) | 0.72 (0.71 to 0.73) |

*15 infants were missing data on clinical risk factors and therefore could not be assigned a clinical risk factor score.
—, calculation cannot be performed because of zero in denominator.

Table 4 Likelihood ratios and post-test probabilities for specific clinical risk score levels and pre-discharge TSB risk zones

| | Post-discharge TSB >95th centile | Post-discharge TSB ≤95th centile | Total | Likelihood ratio | Post-test probability |
|------------------------------|-------------------------------------|-------------------------------------|-------|---------------------|--------------------------|
| Clinical risk score | | | | | |
| 0–7 | 2 | 105 | 107 | 0.15 | 0.02 |
| 8–11 | 12 | 218 | 230 | 0.45 | 0.05 |
| 12–15 | 31 | 272 | 303 | 0.92 | 0.10 |
| 16–19 | 31 | 127 | 158 | 1.98 | 0.20 |
| 20–23 | 19 | 60 | 79 | 2.57 | 0.24 |
| ≥24 | 2 | 5 | 7 | 3.25 | 0.29 |
| Pre-discharge TSB centile | | | | | |
| 0–40 | 1 | 171 | 172 | 0.05 | 0.006 |
| 41–75 | 10 | 388 | 398 | 0.21 | 0.03 |
| 76–95 | 57 | 216 | 273 | 2.16 | 0.21 |
| >95 | 30 | 26 | 56 | 9.43 | 0.54 |

sensitivity = 0.98), but incorrectly labelled as high risk 87% of infants who did not develop significant hyperbilirubinaemia (that is, specificity = 0.13). More stringent positivity criteria (that is, higher score thresholds) result in improved specificity (fewer false positives), but also in a rapid decline in sensitivity (more false negatives) (table 3). The likelihood ratios for specific clinical risk factor score intervals ranged from 0.15 to 3.25 (table 4). The Hosmer-Lemeshow goodness of fit statistic for the clinical risk factor score was non-significant ($p = 0.47$), suggesting good calibration.

Pre-discharge TSB risk zone

The pre-discharge TSB expressed as a risk zone on the hour specific bilirubin nomogram had good discrimination (0.83; 95% CI 0.80 to 0.86) (fig 2) and generated predicted probabilities ranging from 0.01 to 0.59. Using the 40th centile as a threshold to designate infants at risk of developing significant hyperbilirubinaemia identified all but one infant who went on to develop significant hyperbilirubinaemia (that is, sensitivity = 0.99) (table 3) but incorrectly labelled 79% of infants without significant hyperbilirubinaemia as being at risk for developing the outcome (that is, specificity = 0.21). The number of false positives generated using the 40th centile as the positivity criterion is lower than the number generated using a risk score of 8, but the positive predictive value is not much better (13%). This may be tolerated if the clinician's primary objective is to avoid false negatives (maximise sensitivity and NPV). The likelihood ratios for specific risk zones ranged from 0.05 to 9.43 (table 4). The Hosmer-Lemeshow goodness of fit statistic for the pre-discharge TSB risk zone was non-significant ($p = 0.25$).

DISCUSSION

In this study we found that the pre-discharge TSB risk zone was superior to a clinical risk factor score for assessing risk of developing significant neonatal hyperbilirubinaemia. Neither screening strategy could predict the outcome with ≥ 0.98 sensitivity without compromising specificity (0.13 and 0.21, respectively), but the pre-discharge risk zone approach had better discrimination, and multi-level risk zones provided a wider range of likelihood ratios (and predicted probabilities) for risk stratification than multi-level risk scores. For example, a 37 week gestational age (5 points) infant delivered with vacuum assistance (4 points) and oxytocin augmentation (4 points) weighing 3300 g at birth (6 points) and exclusively breast fed at discharge (5 points) would have a clinical risk score of 24, placing her in the highest risk strata with a 29% predicted probability (PPV) of developing significant hyperbilirubinaemia (a post-discharge TSB

>95th centile). If that same infant had a pre-discharge TSB in the high risk zone (>95th centile) a clinician could actually expect that infant to have a significantly higher probability of developing significant hyperbilirubinaemia after discharge (PPV = 54%).

Our clinical risk factor scoring system is similar to clinical risk factor based prediction rules developed by others. Stevenson and colleagues¹⁶ found end tidal carbon monoxide levels, method of feeding, and birth weight were predictive of TSB ≥ 95 th centile on the hour specific bilirubin nomogram at 96 ± 12 hours of life (c-statistic not reported). Chou *et al* identified maternal race, breast feeding, and GA < 38 weeks as consistent predictors of TSB ≥ 20 mg/dl ($c = 0.79$) and TSB \geq age specific AAP criteria for considering phototherapy ($c = 0.69$).¹⁷ Newman and colleagues' risk index,³ which includes exclusive breast feeding, family history of jaundice in a newborn, neonatal bruising, cephalohaematoma, gender, gestational age, and maternal race and age ≥ 25 , had better discrimination than ours ($c = 0.83$), but was developed using a different study design (nested case-control) to predict a different outcome (absolute TSB > 25 mg/dl (428 mmol/l)). Comparison of the pre-discharge TSB risk zone to a slightly modified version of Newman's risk index¹⁸ (developed to predict a post-discharge TSB > 20 mg/dl (342 mmol/l)) found that the discrimination of the TSB risk zone ($c = 0.79$) was superior to the clinical risk factor score ($c = 0.69$).

Our clinical risk factor score differs from others in that it includes oxytocin exposure, a risk factor for hyperbilirubinaemia first recognised more than 25 years ago.^{19–23} Oxytocin exposure may be a surrogate marker for infant size, cephalopelvic disproportion, and nulliparity—all risk factors for hyperbilirubinaemia—although oxytocin may exert some, as yet uncharacterised, direct effect on neonatal bilirubin metabolism. Our rule also includes vacuum extraction, which may be a surrogate marker for large infant size, but also perhaps is a more reliable indicator of bruising and/or cephalohaematoma (known risk factors for hyperbilirubinaemia) than the detection and documentation of these findings in routine clinical care. Finally, similar to Stevenson and colleagues,¹⁶ we found a direct linear relation between greater birth weight and significant hyperbilirubinaemia. The finding in univariate analyses that being small for gestational age (GA) greatly decreased (OR 0.1; 95% CI 0.02 to 0.9) while being large for GA increased the risk of significant hyperbilirubinaemia (OR 1.6; 95% CI 0.9 to 2.9) suggests that size for GA is also a factor in determining risk. The relation of higher birth weight to hyperbilirubinaemia may be mediated by its association with maternal diabetes and subsequent neonatal polycythaemia, or the increased risk of bruising and cephalohaematoma incurred by larger infants.

What is already known on this topic

- Multiple healthcare quality monitoring organisations have recommended pre-discharge screening of newborns for risk of subsequent hyperbilirubinaemia
- Two screening strategies have been recommended—clinical risk factor assessment and determination of hour specific bilirubin values—but their relative predictive performance is not known

Our study has some limitations. It is possible that the initial restriction of the study sample to infants who participated in the hospital sponsored early discharge follow up programme introduced spectrum bias (a form of selection bias). That is, if infants enrolled in the programme were at higher risk of developing significant hyperbilirubinaemia, then the results of the prediction models evaluated may only be generalisable to higher risk infants. Two factors argue against the presence of spectrum bias: (1) the decision to offer participation in the early discharge follow up programme was based entirely on the duration of hospital stay, regardless of risk of developing hyperbilirubinaemia; and (2) the demographics of the study sample closely resemble those of the base population. The fact that not all study infants had a post-discharge TSB performed may have introduced some verification bias. That is, it may have decreased the number of infants with negative test results (low pre-discharge TSB risk zone or clinical risk score) which would have had the effect of over-estimating the sensitivity and under-estimating the specificity of both the TSB and clinical risk factor based prediction models. However, unlike previous studies,^{3 16 17} we restricted the sample used to derive the clinical risk factor score to infants born during months when >75% of infants had post-discharge bilirubin measurements, which should have minimised verification bias. Finally, the paucity of Asians in our study population did not allow discernment of the known association between this ethnic group and the development of significant hyperbilirubinaemia.

In employing either bilirubin screening or risk factor assessment, clinicians must take efforts to avoid measurement error that can affect predictive performance. Methods to improve bilirubin measurement accuracy and inter-laboratory variability include routine calibration of instruments and strict adherence to quality assurance procedures.^{24 25} Providers using clinical risk factor prediction rules to assess risk must be sure to collect accurate information about clinical risk factors using the same definitions as those used in developing the prediction rules. Deviations from those original definitions will compromise predictive accuracy. For this reason, risk assessment strategies that utilise objective factors (such as use of oxytocin or vacuum extraction) will likely produce more generalisable results compared with those that utilise observer dependent factors (such as bruising, cephalohaematoma, and extent of jaundice) and factors that are ambiguous or difficult to define (such as maternal or infant race).

Finally, clinicians should not rely exclusively on the predictions provided by risk assessment strategies at the time of discharge. Events occurring after the prediction has been made, such as inadequate feeding or haemolysis due to cephalohaematoma, may increase the risk of severe hyperbilirubinaemia and will not be reflected in the pre-discharge risk assessment. For example, the one infant in this study whose TSB jumped from the low risk zone prior to discharge to the high risk zone after discharge was breast fed and had his pre-discharge bilirubin measured at 14 hours of life,

What this study adds

- This study shows that the hour specific bilirubin expressed as a risk zone on a bilirubin nomogram is more accurate than a clinical risk factor scoring system for assessing risk of significant hyperbilirubinaemia

which was probably too early to reflect all the factors that result in significant hyperbilirubinaemia, most notably breast feeding problems. For this reason all infants discharged prior to 72 hours should be examined by a health care professional in the first few days after discharge to assess infant well being and the presence of jaundice. The predicted probabilities generated by both risk assessment strategies can be used to determine the timing and frequency of follow up visits for these infants.² For example, infants in the higher risk zones (>75th centile), 26% of whom developed significant hyperbilirubinaemia in the this study, need to be seen closer to 72 hours after birth, while those in the lower risk zones (<75th centile), 2% of whom developed significant hyperbilirubinaemia, could probably be seen closer to 120 hours, unless other risk factors and feeding issues necessitate earlier post-discharge follow up. If follow up within 12–24 hours cannot be assured, infants in the highest risk zone (>95th centile), who had a 54% rate of significant hyperbilirubinaemia, should remain in the hospital until their bilirubin trajectory is elucidated.

ACKNOWLEDGEMENTS

We thank Lois Johnson, MD for sharing data on the bilirubin nomogram and for useful advice on the study design, Emidio Sivieri, MS for his work on the research database, Jennifer Baldwin and Michele Tereschuk for their assistance with medical record reviews and Chris Feudtner, MD, MPH, PhD, Harold Sox, MD, MPH, and Tracy Lieu, MD, MPH for their comments on earlier versions of the manuscript.

Authors' affiliations

R Keren, S Nihtianova, Division of General Pediatrics, Pediatric Generalist Research Group, The Children's Hospital of Philadelphia, PA, USA

A Cnaan, X Luan, Division of Biostatistics and Epidemiology, The Children's Hospital of Philadelphia, PA, USA

V K Bhutani, Department of Pediatrics, University of Pennsylvania, Philadelphia, PA, USA

J S Schwartz, Department of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Funding: Dr Keren was supported by grant number K23 HD043179 from the National Institute of Child Health and Human Development, Bethesda, MD, USA

Competing interests: none declared

REFERENCES

- 1 **Joint Commission on Accreditation of Healthcare Organizations**. Kernicterus threatens healthy newborns. http://www.jcaho.org/about-us/news+letters/sentinel+event+alert/sea_18.htm (accessed 25 August 2004).
- 2 **AAP Subcommittee on Neonatal Hyperbilirubinemia**. Management of hyperbilirubinemia in the newborn infant 35 or more weeks of gestation. *Pediatrics* 2004;**114**:297–316.
- 3 **Newman TB**, Xiong B, Gonzales VM, et al. Prediction and prevention of extreme neonatal hyperbilirubinemia in a mature health maintenance organization. *Arch Pediatr Adolesc Med* 2000;**154**:1140–7.
- 4 **Bhutani VK**, Johnson L, Sivieri EM. Predictive ability of a predischarge hour-specific serum bilirubin for subsequent significant hyperbilirubinemia in healthy term and near-term newborns. *Pediatrics* 1999;**103**:6–14.
- 5 **Maisels MJ**, Gifford K. Normal serum bilirubin levels in the newborn and the effect of breast-feeding. *Pediatrics* 1986;**78**:837–43.
- 6 **Gale R**, Seidman DS, Dollberg S, et al. Epidemiology of neonatal jaundice in the Jerusalem population. *J Pediatr Gastroenterol Nutr* 1990;**10**:82–6.

- 7 Linn S, Schoenbaum SC, Monson RR, *et al.* Epidemiology of neonatal hyperbilirubinemia. *Pediatrics* 1985;**75**:770–4.
- 8 Newman TB, Escobar GJ, Gonzales VM, *et al.* Frequency of neonatal bilirubin testing and hyperbilirubinemia in a large health maintenance organization. *Pediatrics* 1999;**104**(5 pt 2):1198–203.
- 9 Moyer VA, Ahn C, Sneed S. Accuracy of clinical judgment in neonatal jaundice. *Arch Pediatr Adolesc Med* 2000;**154**:391–4.
- 10 Furnival GM, Wilson RW. Regressions by leaps and bounds. *Technometrics* 1974;**16**:499–511.
- 11 SAS. SAS/STAT user's guide, version 8. Cary, NC, 1999.
- 12 Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;**15**:361–87.
- 13 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29–36.
- 14 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;**44**:837–45.
- 15 Hosmer DW, Lemeshow S. *Applied logistic regression*. New York: Wiley and Sons, Inc, 1989.
- 16 Stevenson DK, Fanaroff AA, Maisels MJ, *et al.* Prediction of hyperbilirubinemia in near-term and term infants. *Pediatrics* 2001;**108**:31–9.
- 17 Chou SC, Palmer RH, Ezhuthachan S, *et al.* Management of hyperbilirubinemia in newborns: measuring performance by using a benchmarking model. *Pediatrics* 2003;**112**(6 pt 1):1264–73.
- 18 Newman TB, Escobar GJ, Liljestrand P. Combining laboratory and clinical data for prediction of hyperbilirubinemia in newborns. *Pediatr Res* 2003;**53**:2259.
- 19 Chew WC. Neonatal hyperbilirubinemia: a comparison between prostaglandin E2 and oxytocin inductions. *BMJ* 1977;**2**:679–80.
- 20 Chew WC, Swann IL. Influence of simultaneous low amniotomy and oxytocin infusion and other maternal factors on neonatal jaundice: a prospective study. *BMJ* 1977;**1**:72–3.
- 21 Davies DP, Gomersall R, Robertson R, *et al.* Neonatal jaundice and maternal oxytocin infusion. *BMJ* 1973;**3**:476–7.
- 22 Calder AA, Ounsted MK, Moar VA, *et al.* Increased bilirubin levels in neonates after induction of labour by intravenous prostaglandin E2 or oxytocin. *Lancet* 1974;**2**:1339–42.
- 23 Wood B, Culley P, Roginski C, *et al.* Factors affecting neonatal jaundice. *Arch Dis Child* 1979;**54**:111–15.
- 24 Lo SF, Doumas BT, Ashwood ER. Performance of bilirubin determinations in US laboratories—revisited. *Clin Chem* 2004;**50**:190–4.
- 25 Vreman HJ, Verter J, Oh W, *et al.* Interlaboratory variability of bilirubin measurements. *Clin Chem* 1996;**42**(6 pt 1):869–73.
- 26 Zhang J, Bowes WA Jr. Birth-weight-for-gestational-age patterns by race, sex, and parity in the United States population. *Obstet Gynecol* 1995;**86**:200–8.

Working in paediatrics and not a member of the Royal College of Paediatrics and Child Health?

Have you thought about the benefits of joining?

Benefits and entitlements of membership

- Regular mailings including quarterly newsletter
- Listed in handbook plus your own copy at no extra charge
- Representation on Council (the College's governing body)
- Local support
- *Archives of Disease in Childhood* (BMJ cost to non members £206.00)
- Discount at Spring meeting
- Book club with discounts on books related to paediatrics
- CPD monitoring (fellows and associates)
- Voting at general meetings (fellows and ordinary members)
- Membership certificate (fellows and ordinary members)
- HST monitoring (ordinary members)
- Designation MRCPCH (ordinary members)
- Designation FRCPCH (fellows)

Subscriptions (inclusive of Archives)

| | |
|----------|---------|
| Junior | £140.00 |
| Ordinary | £325.00 |
| Fellow | £380.00 |

Associate £238.00

Concessionary subscriptions are available for those working part time or resident overseas.

Application forms and information on eligibility are available via www.rcpch.ac.uk or by contacting The Membership Section, Royal College of Paediatrics and Child Health, 50 Hallam Street, London W1W 6DE; tel (020) 7307 5619/20/23; fax (020) 7307 5601; email: membership@rcpch.ac.uk.